Determining the structure of lexical entries and grammatical constructions in Construction Grammar*

Hans C. Boas
The University of Texas at Austin

Research in Construction Grammar assumes no strict separation between syntax and the lexicon. However, recent work by Goldberg (1995, 2006) shows that there is indeed a separation between lexical entries and grammatical constructions, including constraints regulating the fusion of grammatical constructions with verbs. This paper argues that Goldberg's characterization of the interactions between lexical entries and grammatical constructions faces some of the same difficulties as the interactions between lexical entries and transformational rules in the Chomskyan framework (Chomsky, 1965, 1981, 1995). Drawing on a variety of corpus data this paper presents specific proposals that should be considered in order to arrive at a solution that overcomes difficulties inherent to Goldberg's approach. Based on a discussion of the concepts of analogy, collocational restrictions, frequency, and productivity this paper proposes to encode different types of semantic, pragmatic, and syntactic information in such a way that it is possible to account for a given utterance from a comprehension perspective, as well as a production perspective.

Keywords: Construction Grammar, grammatical construction, lexical entry, productivity, collocational restriction, frequency

1. Introduction

Research in Construction Grammar (henceforth CxG) takes as its starting point the hypothesis that there is no strict division between what has been called "the lexicon" and "syntax" in other theoretical frameworks (Bloomfield, 1933; Chomsky, 1965, 1981, 1995). For example, Goldberg (1995, p. 7) argues that "both lexical and syntactic constructions are essentially the same type of declaratively represented data structure: both pair form with meaning." On this view, "there are basic commonalities between the two types of constructions (…) that blur the boundary."

Annual Review of Cognitive Linguistics 6 (2008), **113–144**. DOI 10.1075/arcl.6.06boa
ISSN 1572–0268 / E-ISSN 1572–0276

One major advantage of the constructional approach over other syntactic theories is that it requires no separate linguistic "modules," underlying syntactic forms or transformations. There are different strands of CxG, each differing in how they employ formalisms, how they regard psychological plausibility, and how they view the role of motivation in language, among other things. For a comparison of constructional approaches please see Fried & Östman (2004) and Goldberg (2006). The main goal of this paper is to determine the success of the implementation of a unified lexicon-syntax continuum in Goldberg's version of CxG (also known as "Cognitive Construction Grammar") and to suggest a number of methodological modifications that will help improve the current architecture of different constructional approaches. Throughout this paper I will use the terms Construction Grammars and constructional approaches to refer to the different versions of Construction Grammar. The term Construction Grammar (or CxG) will be used to refer to Goldberg's (1995/2006) constructional approach.

For example, a sentence such as *He talked himself blue in the face* (Goldberg, 1995, p. 189) involves an end result state which is typically not directly attributable to the meaning of *talk* alone. Goldberg (1995) argues that there is an independently existing resultative construction capable of contributing additional arguments to a verb's meaning. One of the advantages of this approach is that it avoids positing ad hoc verb senses for verbs such as *talk* by attributing the resultative meaning to an independently existing meaningful construction whose meaning is 'X CAUSES Y TO BECOME Z' (Goldberg, 1995, p. 3). While this view of grammatical constructions captures the distribution of numerous resultatives, it is not entirely clear what prevents sentences such as **He spoke himself blue in the face* (Boas, 2003a, p. 105) from being licensed. That is, although *speak* independently encodes a very similar meaning as *talk*, it does not seem to be compatible with the resultative semantics.

This difference illustrates that the semantics of meaningful constructions interact with the semantics of verbs such that a constructional pattern can be licensed in some cases, but not in others that are closely related in meaning. Understanding this phenomenon necessitates taking a close look at the structure of lexical entries and grammatical constructions to determine the factors that influence the licensing of any type of utterance in CxG. The goal of this paper is to demonstrate that research in CxG can solve the problem of how lexical entries and grammatical constructions interact to produce the full range of attested constructs of a language.

The remainder of this paper is structured as follows. Section two serves as a theoretical comparison by providing a summary of how the separation of the lexicon and syntax is modeled in the Chomskyan framework.[1] This overview serves as the basis for our review of the structure of lexical entries and grammatical construc-

tions in section three. Section four discusses a number of problems faced by one of the most prominent constructional approaches, namely Goldberg's (1995/2006) account. It shows that her characterization of the interactions between lexical entries and independently existing meaningful constructions faces some of the same difficulties as the interactions between lexical entries and transformational rules in the Chomskyan framework. Section five suggests a number of strategies that should be considered in order to arrive at a solution that overcomes difficulties inherent to Goldberg's approach as well as to other constructional accounts. The last section summarizes our findings and suggests directions for further research.

2. The status of lexical entries in the Chomskyan framework

Since the beginnings of generative transformational grammar in the 1950s, linguists of different theoretical persuasions have argued for mechanisms that are capable of "generating" an infinite number of sentences. Underlying this line of research is the idea that "[t]he lexicon is really an appendix of the grammar, a list of basic irregularities" (Bloomfield, 1933, p. 274). Based on the assumption that the lexicon is a separate component of the language faculty, adherents of the Chomskyan framework put forward a number of hypotheses that make it possible to determine "the fundamental underlying properties of successful grammars" (Chomsky, 1957, p. 1). In one of the more recent versions, the *Principles and Parameters* framework (Chomsky, 1981), the lexicon is part of a modular architecture of grammar, which has various levels of representation. Van Riemsdijk & Williams (1986) describe this view of grammar as follows:

> "The system has four levels of representation (D-Structure, S-Structure, LF, and PF) and three rule systems that relate these levels ("Move α"; LF-Movement, control theory, Reconstruction; deletion, filters, phonological rules). In addition, it has a number of separate modules that act like conditions on rule application (as in the case of Subjacency) or like well-formedness conditions on representations (as in the case of most other modules) or on rules (as in the case of X'-theory)." (Van Riemsdijk & Williams, 1986, p. 309)

Without going into the details of how sentential structures are licensed in frameworks like this, a lexical entry for a verb like *love* in (1) would have to specify that it subcategorizes for a direct object NP.

(1) *love*: [NP, NP]

(2) Sascha loves Nicole.

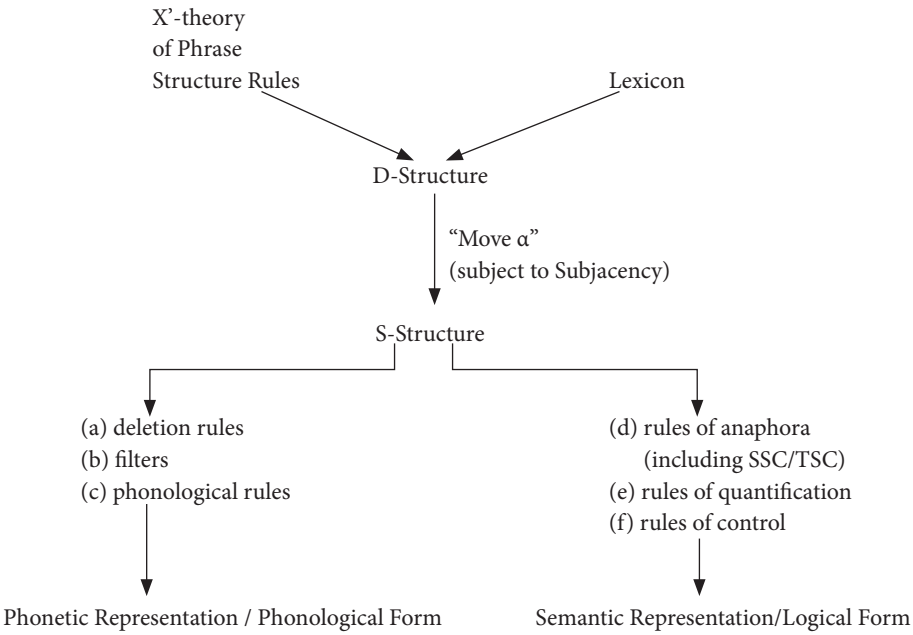

Figure 1. T-model of grammar
(Cf. Van Riemsdijk & Williams, 1986, p. 173)

Once the two NPs have been inserted into the open slots provided by *love* at Deep Structure (due to the application of syntactic phrase structure rules),[2] the Deep Structure representation mirrors the semantic relationships between the verb's two arguments *Sascha* and *Nicole* as in (2). Based on this representation, a number of movement rules and interpretive principles may apply to derive different types of S(urface)-Structures such as *Does Sascha love Nicole?* or *Nicole, Sascha loves.* The application of movement rules is restricted by a number of so-called well-formedness conditions such as the Theta Criterion (Chomsky, 1981) or the Case Filter (*ibid*) that keep such rules from licensing unacceptable sentences. Finally, the surface structure string gets interpreted by the Logical Form (LF) module and spelled out by the Phonetic Form (PF) module.

While the interaction between different linguistic modules has been one of the cornerstones of the Chomskyan framework's architecture for the last four decades, a number of its underlying theoretical assumptions are problematic. The first assumption is that there is a strict separation between the lexicon and syntax. However, it is not clear how it is possible to decide whether semi-productive patterns such as idiomatic constructions and other multi-word expressions should be listed in what is called "the lexicon" or whether they should be captured by syntactic rules, i.e., generated by the "syntactic module" (see, e.g. Fillmore *et al.*, 1988 and

Kay & Fillmore, 1999). This difficulty illustrates that although the lexicon is often thought of as consisting of a theoretically uninteresting repository of idiosyncrasies (Di Sciullo & Williams, 1987), determining its components is not an easy task. For example, is it sufficient to include for the verb *laugh* purely syntactic information about the fact that it is intransitive? What about sentences such as *Joe laughed his head off*? Do we need general syntactic rules that derive this type of idiomatic expression from the verb *laugh*? If there is such a set of rules, how can we prevent them from applying to verbs like *speak*, thereby generating sentences such as **Joe spoke his head off*? These issues suggest that the types of information captured by lexical entries in the Chomskyan framework are not sufficient (see also Sampson 2001, pp. 122–179).

The desire to achieve uniformity of phrase structure configurations has led to an increased number of categories and constraints necessary to restrict interactions between different linguistic modules (see, e.g., Ackerman & Webelhuth, 1998, pp. 30–31). That is, each time a rule is postulated to derive a specific syntactic structure, new constraints must be introduced in order to prevent the rule from over-generating. This problem is shared by other top-down approaches as well, because they typically start their analyses with a small set of sample data. Once more data are analyzed the number of existing constraints does not suffice to restrict the generative mechanisms from over-generation. Finally, it is not clear how much information should be contained in a lexical entry in order for higher-level generative mechanisms to produce the full range of attested utterances while ruling out unattested ones.[3] In the following section I turn to the architecture of CxG in order to see how this alternative framework models the interaction between lexical and syntactic information.

3. The role of lexical entries in Construction Grammar

All constructional approaches share the idea that there are no theoretical distinctions between different areas of grammar such as core and periphery. Arriving at an adequate theory of language mandates examining more than a select number of syntactic phenomena, that is, the aim is to "undertake a commitment in principle to account for the entirety of each language." (Kay & Fillmore, 1999, p. 1) In contrast to most other theories of language, constructional approaches are laid out to be a non-derivational and non-modular theory of language without any strict division between the lexicon and syntax. One difference between the various constructional approaches is that Fillmore *et al.* (1988) and Kay & Fillmore (1999) — in contrast to Goldberg (1995/2006) and Croft (2001) — do not regard all grammatical constructions as meaningful or motivated. On Goldberg's view of

grammar, "much of language is idiosyncratic to varying degrees and must therefore be learned." (Goldberg, 1999) At the center of Goldberg's approach to CxG are "constructions" which are pairings of forms with meanings. For example, Goldberg (1995) defines a construction as follows:

> C is a CONSTRUCTION iff$_{def}$ C is a form-meaning pair <F_i, S_i> such that some aspect of F_i or some aspect of S_i is not strictly predictable from C's component parts or from other previously established constructions. (Goldberg, 1995, p. 4)

In more recent work, Goldberg (2006, pp. 5–9) modifies her definition of constructions by making room for pragmatic and discourse information under the pole of the construction (for alternative definitions of constructions, see Croft (2001, pp. 17–21) and Fried and Östman (2004, pp. 18–23)). In contrast to other theoretical approaches such as the Chomskyan framework, which claim that only words encode meanings, Goldberg (1995) proposes that both words and constructions convey contentful meaning in the interpretation of sentences. In order to illustrate the mechanisms underlying the formation of sentences, it is necessary to first take a look at the structure of lexical entries in CxG. Then, we see how lexical entries interact with grammatical constructions.

3.1 The structure of lexical entries

Adapting Fillmore's (1982, 1985) theory of Frame Semantics and Lakoff's (1987) Idealized Cognitive Models, Goldberg (1995, pp. 24–66) describes lexical entries relative to some particular background frame that designates an idealization of a "coherent individuatable perception, memory, experience, action, or object" (Fillmore, 1977, p. 84). To illustrate, compare the lexical entry for the verb *talk* in (3).

(3) talk < **talker** > (Goldberg, 1995, p. 189)

The lexical entry for the verb *talk* includes information about one participant-role (the talker) which is the crucial part of a verb's frame semantics. The bold face notation indicates that the talker is a profiled role. "Lexically profiled roles are entities in the frame semantics associated with the verb that are obligatorily accessed and function as focal points within the scene, achieving a special degree of prominence" (Goldberg, 1995, p. 44). According to Goldberg (1995, p. 28), lexical entries of the type in (3) only "make reference to world and cultural knowledge" but do not need to include syntactic information as "the mapping between semantics and syntax is done via constructions, not via lexical entries." To see how this mapping is achieved in Goldberg's framework, we now turn to the status of grammatical constructions and their interactions with lexical entries.

3.2 Interaction between verbs and constructions

According to Goldberg, constructions are meaningful entities that pair form with meaning independently of the particular verbs that instantiate them. In other words, constructions are capable of conveying meaning in the interpretation of sentences because they are associated with their own specific semantics by bearing their own arguments. The meaning of a construction is encoded in terms of relationships holding between its abstract semantic roles, also known as argument roles (representing event types). To illustrate, consider the resultative construction in Figure 2, which pairs a specific meaning 'X CAUSES Y TO BECOME Z' with a particular form, namely 'Subj V Obj Xcomp.' (Goldberg, 1995, p. 3)

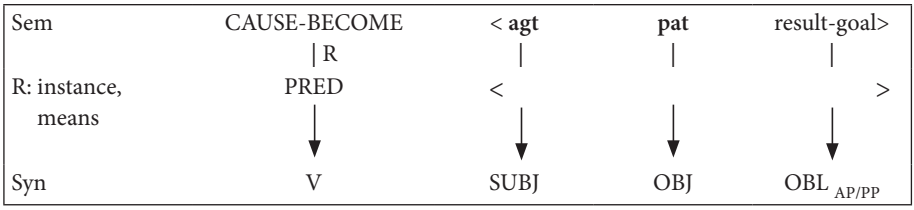

Figure 2. Resultative Construction
(Goldberg, 1995, p. 189)

The top line in the boxed diagram in Figure 2 captures the semantics of the resultative construction in terms of three semantic roles (agent, patient, and result-goal) and the CAUSE-BECOME relation. When a lexical entry of a verb fuses with a construction, the verb's participant roles get integrated into the construction (see the open array to the right of PRED in the middle line in Figure 2). For example, when *talk* fuses with the resultative construction in Figure 2, the verb provides its participant role, namely the talker.[4] Since the verb's talker role can be construed as an instance of the construction's agent role, the two roles are compatible, i.e. the semantics of the verb and the semantics of the construction may fuse. Given the compatibility between *talk* and the resultative construction, Goldberg claims that the construction may also contribute its own patient and result-goal arguments to the predicate's role array. As a result of the verb's fusion with the construction we see how the three semantic arguments are realized syntactically in the bottom line in Figure 2. The agent is realized as the subject, the patient is realized as the direct object, and the result-goal argument is realized as an oblique object. The mechanisms underlying the fusion of verbal and constructional semantics are the same for other types of constructions in Goldberg's framework, such as the caused-motion construction, the ditransitive construction, and the *way*-construction, among others.

Goldberg's approach has a number of advantages over other frameworks. Since the patient and result-goal arguments are arguments of the construction

and not arguments of the verb *talk*, "the verb retains its intrinsic semantic representation, while being integrated with the meaning directly associated with the construction." (Goldberg, 1995, p. 189) This means that it is possible to avoid "the claim that the syntax and semantics of the clause is projected exclusively from the specifications of the main verb." (Goldberg, 1995, p. 224) In other words, under Goldberg's approach the number of lexical entries is kept to a sizable number by excluding specific lexical entries such as one for *talk* that would exclusively license all arguments in sentences such as *He talked himself blue in the face*. (Goldberg, 1995, p. 189) Another benefit of the constructional approach is that it is possible to factor out the syntactically relevant aspects of sentence meaning by attributing them to the semantics of independently existing meaningful constructions (see Goldberg, 1995. pp. 28–30). With this overview, we now turn to the question of whether the constructional approach is capable of overcoming the problems associated with the interaction between lexical entries and the various rule types of the modules in the Chomskyan framework (see section two above).

4. Problems with the interaction between verbs and constructions

While Goldberg's approach to argument structure offers effective solutions to many problems of analyses couched in other theoretical frameworks (see Goldberg, 1995, pp. 7–23, 101–108), it, too, faces a number of difficulties when trying to account for a larger range of data. As the following sections illustrate, many of the problems involve the interaction between lexical entries and grammatical constructions.

4.1 Delimiting the Power of Constructions

Previous work by Kay (1996, 2002b), Nemoto (1998), Iwata (1998) and Boas (2002b, 2003a) has pointed out that it is not always clear how the fusion of verbal semantics and constructional semantics can be constrained. For example, Boas (2003a) argues that under a Goldberg-type approach to CxG, lexical entries for communication verbs share a very similar structure as in the following examples.

(4) a. talk < **talker** > (Goldberg, 1995, p. 189)
b. speak < **speaker** >
c. whisper < **whisperer** >
d. grumble < **grumbler** >
e. murmur < **murmurer** >
f. sigh < **sigher** > (cf. Boas, 2003a, p. 106)

Lexical entries such as those in (4a)–(4f) encode semantic information about the agent of the respective verbs. As such, they are all similarly capable of licensing regular declarative sentences such as those in (5a)–(5f).

(5) a. Miriam talked (to Joe).
b. Miriam spoke (to Joe).
c. Miriam whispered (to Joe).
d. Miriam grumbled (to Joe).
e. Miriam murmured (to Joe).
f. Miriam sighed (to Joe).

Since the lexical entries in (4) are all capable of licensing the respective sentences in (5), one would expect that the types of lexical information provided by these lexical entries are also sufficient for the licensing of other types of constructions. However, as Boas (2003a, pp. 105–107) points out, this is not the case. Compare the following sentences.

(6) a. Miriam talked herself blue in the face.
b. *Miriam spoke herself blue in the face.
c. ?Miriam whispered herself blue in the face.
d. *Miriam grumbled herself blue in the face.
e. *Miriam murmured herself blue in the face.
f. ?Miriam sighed herself blue in the face. (cf. Boas, 2003a, pp. 105)

These examples illustrate that the six semantically related verbs do not exhibit a uniform distribution in resultative constructions. What could be the reason for this discrepancy? In section three we saw that *talk* is capable of fusing with the resultative construction because its talker role can be construed as an instance of the construction's agent role, which in turn means that the two roles are compatible. This allows the verb *talk* to fuse with the resultative construction, which in turn is capable of supplying its two argument roles (patient and result-goal), thereby leading to a resultative interpretation. Due to the distribution of the six verbs outside of resultative constructions, one might expect them to exhibit a similar distribution when it comes to resultatives.

However, based on the principles advocated by Goldberg for the fusion of verbal and constructional semantics, it is not clear what factors may stop the resultative construction from fusing with the semantics of *speak*, *whisper*, and *grumble* to yield a resultative interpretation. In other words, what are the circumstances under which the integration of the verbal semantics in (4) into the resultative construction are blocked? Since none of Goldberg's construction-specific constraints are capable of ruling out the unattested sentences in (6), it is necessary to consider alternative methods for blocking the integration of verbs into the resultative

construction. Note that the varied distribution among semantically related verbs is not limited to resultatives: it is also found with other constructions such as the *way*-construction (Goldberg, 1995, pp. 199–218) in (7) and the metaphorical extensions of the ditransitive construction (Goldberg, 1995, pp. 148–150) in (8).

(7) a. Miriam talked her way into the office.
b. *Miriam spoke her way into the office.
c. ?Miriam whispered her way into the office.
d. ?Miriam grumbled her way into the office.
e. *Miriam murmured her way into the office.
f. Miriam sighed her way into the office.

(8) a. *Miriam talked Joe a fairy tale.
b. *Miriam spoke Joe a fairy tale.
c. Miriam whispered Joe a fairy tale.
d. *Miriam grumbled Joe a fairy tale.
e. ?Miriam murmured Joe a fairy tale.
f. ?Miriam sighed Joe a fairy tale.

These sentences illustrate that although the verbs are semantically related, some may unify with the *way*-construction or the ditransitive construction, while others may not. We return to a discussion of how to block verbs from unifying with constructions in Section 5.4.

4.2 Determining the range of postverbal arguments

Similar difficulties with the interaction between verbs and constructions arise when it comes to delimiting the semantic range of postverbal arguments. For example, Boas (2003a, pp. 113–116) demonstrates that on the basis of lexical entries of the sort in (9) it is difficult if not impossible to predict the range of argument expressions that may occur with different "verbs of ingesting" (Levin, 1993, pp. 213–217).

(9) eat: < **eater** eaten >

(10) a. Pat ate.
b. Pat ate his food.
c. Pat ate his food up.
d. Pat at his plate clean.

(11) a. Pat chewed.
b. Pat chewed his food.
c. Pat chewed his food up.
d. *Pat chewed his plate clean.

(12) a. *Pat devoured.
b. Pat devoured his food.
c. *Pat devoured his food up.
d. *Pat devoured his plate clean.

(13) a. Pat swallowed.
b. Pat swallowed his food.
c. ?Pat swallowed his food up.
d. *Pat swallowed his plate clean. (cf. Boas, 2003a, p. 114)

The examples illustrate that it is difficult for an independently existing meaningful construction to determine to which verbs it should contribute what types of additional argument roles. This difficulty is not only found in the semantic domain but also in the syntactic domain where phrases closely related in meaning show different types of distribution. The following examples from Boas (2003a) based on the British National Corpus (BNC) show that V-*dead* and V-*to death* exhibit different distributions.

Table 1. Distribution of *dead* in resultative constructions in the BNC

Verb	No. of occur.
Shoot	408
Kill	9
Strike	8
Make, knock	3
Flatten, kick, smite	1

(Boas, 2003a, pp. 130)

The distribution of *dead* and *to death* in Tables 1 and 2 illustrates that verbs differ with respect to their collocational restrictions. Given the current architecture of lexical entries and their interaction with meaningful constructions in a Goldberg-style Construction Grammar, it is hard to predict with what type of resultative phrase a given verb will occur because a construction has no way of distinguishing between the two types of resultative phrases (for similar results, see Verspoor, 1997). Similar problems arise when predictions are made about a verb's alternation patterns based on its membership in a semantic class, an issue to which we now turn.

4.3 The relevance of semantic classes

According to Goldberg (1995, pp. 175–79), the load/spray alternation is licensed by the caused-motion construction and the causative-plus-*with*-adjunct construction. Verbs such as *slather*, *smear*, *brush*, *dab*, and *daub* belonging to Pinker's (1989)

Table 2. Distribution of *to death* in resultative construction in the BNC

Verb	No. of occur.
Stab	114
Beat	74
Batter	39
Frighten	34
Crush	25
Scare	24
Burn	18
Torture	16
Drink, starve	15
Bludgeon, hack	12
Shoot, kick	11
Club	9
Bore, knife, choke	8
Blast, trample, work, worry	7
Love	6
Strangle	4
Dash, poison, kiss	3
Ax, bayonet, boil, bring, clap, suffocate, kick, freeze, spear, spray, stone, suck, gun, hammer, hug, knock, nag, peck, play, rape, shag, sting	2
Annoy, eat, bleed, blend, bug, bully, flog, frit, cudgel, curse, dance, feed, gas, flog, jog, laugh, pitchfork pound, run, schmaltz, scorch, scratch, seduce, shock, sing, smother, squash, squeeze, stamp, strike, suffocate, sweat, whip	1

(Boas, 2003a, p. 131)

slather-class are said to all describe simultaneous forceful contact and motion of a mass against a surface. Based on lexical entries such as the one for *slather* in (14), verbs belonging to the *slather*-class are assumed to exhibit similar distributions of arguments in the load/spray alternation in (15). However, this prediction does not hold as the examples with *brush* in (16) illustrate.

(14) slather < **slatherer**, **thick-mass**, **target** > (Goldberg, 1995, p. 176)

(15) a. Sam slathered shaving cream onto his face.
b. Sam slathered his face with shaving cream.
c. * Sam slathered shaving cream.
d. * Sam slathered his face.
e. * Shaving cream slathered onto his face. (Goldberg, 1995, p. 176)

(16) a. Joe brushed tooth paste onto his teeth.
b. Joe brushed his teeth with tooth paste.

c. *Joe brushed tooth paste.
d. Joe brushed his teeth.
e. *Tooth paste brushed onto his teeth. (Boas, 2003b, p. 32)

This problem illustrates that it is difficult to predict a verb's range of arguments based on semantic class membership. Whereas Pinker (1989), Levin (1993), and Goldberg (1995) seek a way of determining the range of a verb's syntactic arguments based on the membership of a verb in a specific semantic class, the examples in (14)–(16) suggest that semantic classes will have to be defined more precisely. Once this important step is accomplished, it may be possible to accurately determine a verb's range of argument based on its semantic class membership. See Boas (2006) and Ruiz de Mendoza Ibáñez & Mairal (2007, 2008) for proposals along these lines.

4.4 Lexical entries, constructions, and the architecture of Construction Grammar

Although CxG has a number of empirical and theoretical advantages over other frameworks, the difficulties discussed in the previous sections illustrate that a Goldberg-style CxG is also problematic. Following previous work by Kay (1996, 2002b), Nemoto (1998), and Boas (2003ab), I propose that the problems discussed above are due to the fact that the interactions between lexical entries and meaningful grammatical constructions are difficult to constrain. More specifically, I suggest that although most work in CxG explicitly denies any strict separation between the lexicon and syntax, such a split, although subtle, does indeed exist. Consider the interaction between lexical entries and constructions in the following figure.

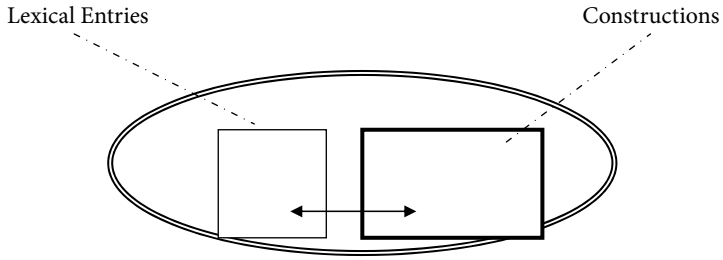

Figure 3. Interaction between Lexical Entries and Constructions

Figure 3 is meant to represent a speaker's linguistic knowledge, including knowledge about lexical items and grammatical constructions. On the Goldbergian view of CxG, lexical entries and constructions interact by fusing together, thereby licensing sentences.[5] Contrasting this constructional approach with the licensing

of sentences in the Chomskyan framework (see section two), certain parallels emerge. That is, in both approaches the amount of information contained in a lexical entry seems to be kept to a minimum.[6] In order to license sentences, a rule or a construction interacts with the information contained in a lexical entry. It is this interaction that has been shown to be difficult to constrain in both the Chomskyan framework and Goldberg's CxG approach.

To solve this problem, one could argue with Goldberg (1995, p. 43) that "[p]articipant roles are instances of the more general argument roles and capture specific selectional restrictions as well." However, this is not the case as we have seen in Section 4.1–4.2. There, I demonstrated that although the verbal semantics are claimed by Goldberg to be instantiations of constructional semantics, it is still necessary to include more specific semantic and syntactic information in a verb's lexical entry in order to be able to predict its distribution with a variety of different constructions. As such, Goldberg's present selection restrictions are not specific enough (cf. (4)–(16)).

Our discussion illustrates that unless viable solutions are found to the problems surrounding the interaction of lexical entries and grammatical constructions, research in CxG may face problems similar to those encountered by analyses in the Chomskyan framework that attempt to delimit the application of generative mechanisms (PS-rules, transformations, etc.) by formulating an ever-increasing number of constraints and categories. To avoid these difficulties, research in CxG first needs to determine the range of a construction's application by following a bottom-up approach to linguistic description (see Boas, 2002ab, 2003ab). Only when we know more details about the distribution of a given constructional pattern will we be able to better understand how lexical entries and constructions interact to license the utterances of a language. Then we will be in a good position to solve the types of problems pointed out in Sections 4.1–4.3 above (selection restrictions, collocational restrictions, defining proper semantic classes for effective linking mechanisms). The following sections present in more detail a number of proposals, which I believe are crucial for resolving the problems currently arising from the interactions between lexical entries and constructions in CxG.

5. Unifying the structures of lexical entries and grammatical constructions

5.1 The need for a corpus-based bottom-up approach

The first proposal is a methodological one. That is, in order to find out the range of a construction's distribution it is first necessary to determine to the greatest

extent possible what verbs may occur with different argument structures licensed by a construction. For example, in previous work (Boas, 2002b, 2003ab) I demonstrated that resultative and caused-motion constructions are much more restricted in their range of application than previously assumed. Based on more than 6000 resultative and caused-motion constructions in the British National Corpus I showed that the combinations of verbs and their postverbal arguments are syntactically, semantically, and pragmatically highly restricted. This observation led me to argue for the existence of so-called mini-constructions, which represent conventionalized senses of verbs including syntactic, semantic, and pragmatic information. These mini-constructions are in principle comparable to Croft's (2003) verb-class and verb-specific constructions. Mini-constructions are lexicalized representations necessary for predicting the exact distribution of a verb in resultative and caused-motion constructions, including information about the types of resultative phrases that may occur with a given verb. For example, by encoding collocational information about the different types of postverbal arguments that may occur with a verb, it becomes possible to model whether a given verb will occur with *to death* or with *dead* as its resultative phrase (see Tables 1 and 2 above). This alternative analysis demonstrates the general need for introducing a corpus-based bottom-up approach to CxG (see also Boas, 2005, 2007).

Including the widest possible range of corpus data for the description of linguistic phenomena is not a new idea. For example, Hunston & Francis (2000) adopt a corpus-based approach for their description of a lexical grammar of English. On this view, words occur with a number of major "patterns", i.e., "the elements that follow it" (Hunston & Francis, 2000, p. 51). A description of a given pattern involves not only an analysis of a few select words, but also the full range of words that may occur with this pattern. Similarly, adherents of usage-based models (Bybee, 1985, 2001; Langacker, 1987, 2000) have argued for the inclusion of grammatical information in taxonomic hierarchies that also represent redundant information. The idea behind this methodology is that only by determining the full range of lexical items occurring with a certain pattern is it possible to determine the degree of generality of a construction. Research within the Lexicon Grammar framework as developed by Maurice Gross has shown that "[m]ore syntactic properties of sentences than usually thought depend on the main verb. (…) The systematic description of French verbs (or simple sentences) has shown that no two verbs have the same syntactic properties." (Gross, 1994, p. 214) To summarize, adopting a corpus-based bottom-up approach to the description of lexical entries and grammatical constructions is essential to overcoming the problems pointed out in section four above (see also Fillmore *et al.*, 2003).

5.2 The status of abstract meaningful constructions

In addition to determining the interaction between lexical and constructional information, it is necessary to figure out the exact status of abstract meaningful constructions vis-à-vis other types of constructions. Recall that on the Goldbergian view, argument structure constructions are independently existing form-meaning pairings capable of contributing additional arguments to verbs. However, the exact theoretical status of this type of construction is far from clear. One of Goldberg's main motivations for the existence of independently existing meaningful constructions is to avoid implausible verb senses, e.g., an extra sense of *sneeze* to account for sentences such as *Kim sneezed the napkin off the table*.

While abstract argument structure constructions seem to be of great importance for the understanding of novel utterances, they are not sufficient for the production of all novel utterances. As we have seen in section four, verbs exhibit idiosyncratic behaviors when it comes to the distribution of postverbal arguments. This diversity makes it difficult if not impossible to predict a verb's distribution of arguments based on a construction's capability of contributing arguments to the semantics of a verb. With respect to resultative constructions, I have argued for a pattern of analogical association that licenses the production of novel resultatives (see Boas, 2003a, pp. 260–277). In the case of a non-conventionalized resultative like *Kim sneezed the napkin off the table*, the speaker utilizes existing conventionalized mini-constructions and contextual background information. That is, a specific conventionalized sense of *blow* is already associated with the [NP V NP PP] syntactic frame. When there is sufficient semantic and pragmatic overlap between the basic sense of *sneeze* and the conventionalized resultative sense of *blow*, then *sneeze* may be used with the same syntactic frame as *blow*.[7] This alternative to Goldberg's approach shows that it is in principle possible to account for the production of novel utterances based on existing conventionalized knowledge without having to rely on abstract meaningful constructions à la Goldberg (caused-motion, resultative, ditransitive, etc.) (see also Boas, 2005, 2007).[8]

Analyzing different sets of data, Kay (2002b) also finds variations in productivity between various types of patterns, such as the *all*-cleft pattern in (17) and the [A *as* NP] pattern in (18).

(17) a. All that we had to say to them was that we intended to tax them more severely. (BNC)
b. All that one has to do is to start training earlier. (BNC)
c. All I want to get is out of the flat, … (BNC) (Kay, 2002b)

(18) a. dumb as an ox
b. poor as a church mouse

c. green as grass
d. dead as a doornail (Kay, 2002b)

Reviewing the differences between the two patterns exemplified by (17) and (18), Kay (2002b) demonstrates that the *all*-cleft pattern exhibits a regular *wh*-cleft syntax (compare (17b) with *What one has to do is to start training earlier*) and a below-expectation reading (compare (17b) with *#Everything one has to do is to start training earlier*). He points out that "this construction is fully productive, being lexically constrained only with respect to the left isolate constituent of the subject phrase." (Kay, 2002b) In contrast, Kay shows that the [A *as* NP] pattern must be learned individually due to its greater degree of idiosyncrasy. For example, he cites the examples in (19) to demonstrate that [A *as* NP] expressions differ with respect to whether they can occur with comparative morphosyntax or not. In contrast, other [A *as* NP] expressions only occur in the comparative form as in (20), according to Kay.

(19) a. deader than a doornail
b. hotter than hell
c. bigger than a house
d. flatter than a pancake
e. *happier than a lark
f. *quicker than a wink
g. *easier than pie
h. *drier than a bone (Kay, 2002b)

(20) a. larger than life.
b. *large as life
c. better than a jab in the eye with a sharp stick/eating a bug
d. *as good as a jab in the eye with a sharp stick/eating a bug (Kay, 2002b)

Examples such as those in (19) and (20) lead Kay (2002b) to suggest that the [A *as* NP] pattern is not as productive as other types of constructions such as the *all*-cleft construction. In other words, the regular [A *as* NP] pattern must be distinguished from its comparative [A-*er than* NP] counterpart since not all [A *as* NP] patterns have a comparative counterpart. In fact, he argues that a pattern having certain one-shot extensions does not guarantee its full productivity: "The existence of a handful of novel literary or poetic examples in a corpus does not prove that a pattern of coining is a productive construction; it only illustrates the familiar fact that nonce coinages do occur."

Kay's observations about productivity differences point in the same direction as our previous discussion. Given the current architecture of CxG it is not clear what the exact status of independently existing meaningful constructions in

Goldberg's framework is, and whether these constructions are equally necessary for the interpretation and the production of utterances. In order to find a solution to these problems, research in CxG must come to grips with three fundamental issues: measuring productivity, differences in meaning between constructions, and differences between production and comprehension. The last point is especially important because there is evidence which suggests that different types of constructional knowledge are necessary for encoding versus decoding (see Fillmore *et al.*, 1988). For example, Boas (2005) proposes that item-specific knowledge is crucial for correctly encoding resultatives (in terms of mini-constructions), while still acknowledging the existence of higher-level schematic resultative constructions that come into play when decoding resultatives, especial non-conventionalized ones (see also Boas, 2007).

5.2.1 *Measuring productivity*

The first of these issues concerns the criteria that are to be used to decide on the degree of productivity exhibited by a given constructional pattern. Assume that one follows the corpus-based bottom-up approach proposed in the previous sections. Having gathered extensive corpus data on a certain pattern one finds no exceptions or irregularities, such as the *all*-cleft construction or the subject-predicate construction that requires an English verb to agree in number with its subject. In this case one would want to postulate an abstract construction with a narrowly defined range of application without having to state any (or very few) constraints. But what do we do in case a constructional pattern is less productive? Where and how do we encode the relevant constraints? In the case of resultatives, Boas (2003a) argues for encoding the relevant restrictions in mini-constructions that each represent conventionalized senses of verbs. This approach appears to be quite successful in accounting for the distribution of more than 6000 resultatives in the BNC. But while constructions exhibiting full productivity (such as the *all*-cleft construction) and those exhibiting a very limited degree of productivity (such as resultatives) are at opposite ends of the productivity spectrum, we still have no principled way of deciding how to measure the productivity of constructions that fall in between the opposite ends of the productivity spectrum. In other words, it is necessary to determine the factors that should be used in measuring and evaluating the degree of a construction's productivity (this may differ depending on the speaker, the genre, etc.). Further research in CxG must address this issue, in particular if we want to arrive at an understanding of the division of labor between what have been labeled abstract constructions, lexical entries, and constraints that limit the interaction between the two (for a proposal along these lines see Barðdal, to appear).

5.2.2 *Differences in meaning between constructions*

The second issue that needs to be addressed concerns the differences in meaning between various types of constructions. According to Goldberg (1995, p. 224), "meaningful constructions" are "free-standing entities, stored within the lexicon alongside lexical items, idioms, and other constructions that may or may not be partially filled." (1995, p. 221) In the case of constructions such as the ditransitive construction, which pairs a specific meaning ('X CAUSES Y TO RECEIVE Z') with a specific form (Subj V Obj Obj_2), it seems possible to determine the form-meaning pairing in a fairly straightforward way. Based on a number of constraints, it is also possible to determine for most cases whether the ditransitive construction is capable of contributing additional argument roles to a verb's array of participant roles. This process would license sentences such as *Sally baked her sister a cake* (Goldberg, 1995, p. 141), where *bake* does not encode an intended transfer meaning independently of the ditransitive construction.

But what do we do about more abstract constructions such as different types of passive constructions? For example, Ackerman & Webelhuth (1998) identify a total of 14 distinct yet related German passive constructions. Although each construction has its own specific range of application, it is difficult to assign each a precise meaning. Although the meanings of the passives are quite different from their active counterparts, it is unclear whether this difference in meaning is to be attributed to a fusion between verbal and constructional semantics similar to that advocated by Goldberg for meaningful constructions such as the ditransitive or the *way*-constructions. Also, note that passive constructions do not typically contribute additional arguments to the semantics of the verb, but rather "re-arrange" a verb's syntactic frame. In this connection, consider the subject-predicate construction, which is even more abstract than the passive construction. Although we know the subject-predicate construction requires the verb to agree with its subject, we are not clear about its meaning. Again, we are faced with a construction which neither contributes argument roles nor does it seem to encode any meaning. The differences in meaning between the ditransitive, caused-motion, resultative, and *way*-constructions discussed by Goldberg (1995) and more abstract constructions such as the passive and the subject-predicate constructions suggest that constructions differ significantly with respect to their ability to contribute additional arguments to verbs, or other constructional properties. Future research in CxG needs to address these fundamental differences in meaning in order to determine what types of meanings should be encoded at what level of abstraction.

5.2.3 *Production and comprehension*

The third issue that needs clarification is the different role that constructions play in the production and comprehension of utterances. On Goldberg's (1995) view,

constructions allow us "to capture generalizations across instances." This leads her to claim that "what is stored is the knowledge that a particular verb *with its inherent meaning* can be used in a particular construction. This is equivalent to saying that the composite fused structure involving both verb and construction is stored in memory." (1995, p. 140)

Our discussion in section four has shown that her approach towards the interaction of verbal and constructional semantics is problematic when it comes to predicting the range of possible arguments in different argument structure constructions. These difficulties suggest that the composite fused structures involving verbs and constructions are not stored in memory. If this were the case, we would be able to predict the full range of arguments capable of occurring with any given verb because the fused structure would include more detailed information than that provided by Goldberg's current type of lexical entries. I propose that these problems are due to the differences in processing in comprehension and production. That is, upon hearing a novel sentence such as *Joe sneezed the napkin off the table*, it may be possible to identify its meaning based on the syntactic frame and an abstract meaningful construction that links the [NP V NP PP] frame with a caused-motion semantics. However, this abstract construction only exists because it represents a generalization over a large number of individual verbs that are conventionally associated with the caused-motion semantics. In the case of *sneeze*, *blow* is the prototypical air-emission verb that conventionally pairs a caused-motion semantics with a [NP V NP PP] frame (for details, see Boas 2003a, pp. 260–277). In other words, the independent caused-motion construction is an abstraction over individually learned verbs that pair a [NP V NP PP] frame with a caused-motion semantics. It is this generalized constructional schema that allows for a straightforward interpretation of novel utterances (see also Langacker, 2000 and Goldberg, 2006).

The difficulties that arise when one wants to predict the range of a verb's arguments illustrate that an abstract construction does not provide sufficient information in combination with simplified lexical entries. In other words, the division of labor between constructions and lexical entries is not "equivalent to saying that the composite fused structure involving both verb and construction is stored in memory" as advocated by Goldberg (1995, p. 140) This problem suggests that there is a fundamental difference in the role that constructions play in the comprehension and the production of utterances. More specifically, while abstract constructions may be sufficient for comprehension, for production we need to refer to more detailed information in order to arrive at correct predictions about the distribution of a verb's arguments (see also Boas, 2008). Future research in CxG needs to address the question of where this more detailed information should be included, i.e., whether it is possible to arrive at generalized constraints over a narrowly defined class of verbs, or whether — as in the case of resultatives — idiosyncratic

restrictions need to be encoded at the level of individual verb senses. I suspect that the answer to this question will crucially depend on a construction's productivity as well as its inherent meaning and level of abstraction (see following section).

5.3 Significance of frequency data

Another important point relates directly to the status of abstract constructions discussed in the previous section. Adopting ideas from usage-based approaches (e.g., Bybee, 1985, 2001; Langacker, 1987, 2000), Goldberg (1995, p. 134) argues that token frequency and type frequency play a crucial role in the classification of new verbs.[9] She points out that "type frequency of a particular process (or a particular construction) (…) plays a crucial role in determining how likely it is that the process may be extended to new forms: the higher the type frequency, the higher the productivity." (1995, p. 134) Goldberg reviews acquisition data that shed light on how children categorize combinations of linguistic units they have previously not encountered (thereby over-generalizing to a certain degree), and points out that "it is not necessary that each new entry be stored as an additional member of a cluster, throughout the speaker's life." This leads her to propose that it "is possible that once a critical number of instances in a particular cluster is learned — insuring that novel instances that fall into the class will be included — new cases are no longer stored in memory since they would provide only entirely redundant information." (1995, p. 136)

While Goldberg is certainly on the right track with respect to the influence of frequency data on categorization and schematization, she does not provide further clues as to how frequency data might be empirically used to determine a construction's productivity. Including frequency data in CxG is a very important step as various psycholinguistic studies strongly suggest that different patterns of frequency of use are relevant for a speaker's grammatical knowledge and the level of entrenchment of linguistic information (cf. Bybee, 1985; Harris, 1998; Baker, 1999). Different levels of frequency are well attested in corpus-based studies of various phenomena. Compare, for example, the frequency rate of different types of resultative phrases occurring with the "*drive*-mental state" sense of *drive* in the BNC (for details, see Boas, 2003a, p. 186).

Assuming that a theory of grammar models the usage-based knowledge necessary to understand and produce all utterances of a language, it is essential to include in our theory frequency-based information of the type shown in Table 3. For example, as a part of our lexical entry we would want to include information that the resultative phrase collocating with the specific sense of *drive* as in *Michael drove Joe crazy* is realized as an adjectival phrase in 77% of all cases and as a prepositional phrase in 23% of all cases found in the British National Corpus.[10]

Table 3. Frequency of resultative phrases occurring with "*drive*-mental-state" sense of *drive* in the BNC

Resultative Phrase	No. of occur.
Mad / to madness	108 / 5
Crazy	70
To distraction	27
Insane / to insanity	23 / 1
Wild	22
Nuts	18
Up the wall	13
To suicide	9
To despair	8
To desperation	7
Batty	4
Dotty	4
Crackers	4
Into a frenzy	3
Over the edge	3

(cf. Boas, 2003a, p. 186)

Besides including frequency information about a construction's different slots, research in CxG should also address the question of how to encode information about a construction's frequency vis-à-vis other types of constructions. For example, in CxG there is currently no systematic notation for capturing information about the fact that instances of the subject-predicate construction occur with a higher frequency than instances of other types of constructions such as the passive construction or the *way*-construction. Given the important role of frequency in linguistic processing (Bybee, 2001, pp. 19–34), inclusion of frequency data in constructional research will help to determine the organizational relationships between different constructions within a hierarchically organized lexicon-syntax continuum. Such an understanding will also shed light on the relations between a construction's frequency, productivity, and schematicity, as the following figure illustrates.

Figure 4 illustrates the types of results made possible by including a construction's frequency information in a CxG description of the lexicon-syntax continuum. The two double arrows point to opposite ends of two spectrums. The spectrum on the left side encompasses the different levels of abstraction of constructions. For example, the subject-predicate construction is a rather abstract construction as it encodes very little meaning. In contrast, the passive construction and the *way*-construction encode a more specific meaning than the subject-predicate

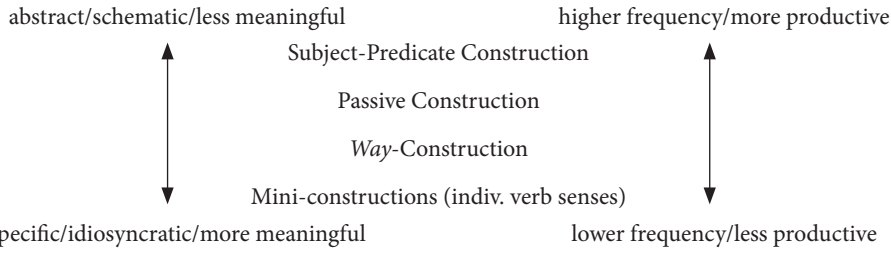

Figure 4. Organization of the lexicon-syntax continuum[11]

construction. Mini-constructions, which are representations of individual verb-senses such as the *drive*-mental-state sense of *drive*, are even more specific than the passive construction or the *way*-construction as they denote a more explicit semantic space that is narrower and more idiosyncratic.

Our discussion of the idiosyncrasy of resultatives has shown that there is a direct link between a construction's idiosyncrasy and both its frequency and its productivity. It is this link that is represented by the spectrum on the right side in Figure 4. That is, mini-constructions that license resultatives are not only extremely idiosyncratic, but they are also very limited in their productivity and frequency vis-à-vis other types of more abstract constructions such as the *way*-construction, the passive construction, or the subject-predicate construction. The latter constructions are more abstract, less constrained, and hence more productive. This observation suggests that a construction's abstractness correlates directly with its frequency and productivity. For a detailed proposal about the relationship between a construction's type frequency, semantic coherence, and productivity see Barðdal (to appear) on the argument structure of novel verbs in Icelandic. Future research in CxG should study this correlation more closely by including a construction's frequency data in its description. Such an approach will not only shed light on the link between a construction's productivity and its abstractness, but will also help to arrive at a better understanding of how the lexicon-syntax continuum is structured.

5.4 The status and structure of lexical entries

The fourth and final proposal concerns the status and structure of lexical entries in CxG. Given the importance of the lexicon-syntax continuum for CxG, the Goldberg-style notation of lexical entries and constructions as discussed in Section 3 should be re-considered for the following reasons. Goldberg uses two different types of notations for lexical entries and constructions. If we take the lexicon-syntax continuum seriously, it is not clear why entities placed at different points along the same continuum should be structurally different. In Section 3 we have

seen that these interactions are difficult to constrain, a feature that is also common to the Chomskyan framework. To avoid these problems, research in CxG should aim for a more unified notation for constructions with different levels of abstractness (cf. the notation in Fillmore & Kay, 1993). Encoding the idiosyncratic form-meaning pairings of individual verb senses and idiomatic constructions using the same notation that is used to describe the properties of more abstract constructions will create a true lexicon-syntax continuum in which there are no boundaries between lexical entries and grammatical constructions.[12] Following Croft (2001), I propose to use a uniform notation representing the link between form and meaning as shown in Figure 5.

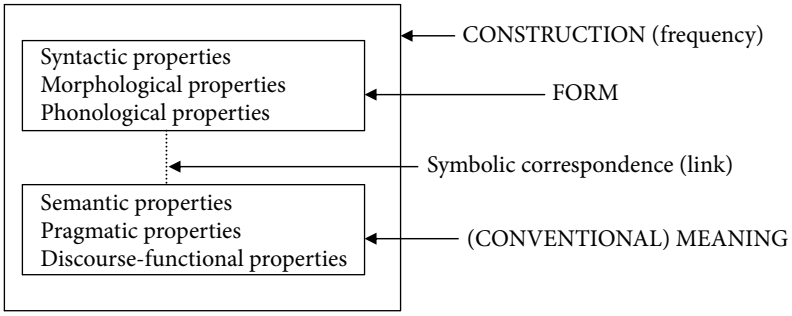

Figure 5. Uniform notation for constructions exhibiting different degrees of abstractness (based on Croft, 2001, p. 18)

Constructions as illustrated in Figure 5 are taken to be symbolic units that link form with meaning (cf. Langacker, 1987, p. 60; Croft, 2001, p. 18). Including information about a construction's frequency in its form component (top box in fig. 5) allows us to directly correlate a construction's frequency with its level of abstraction, which is represented by its meaning component (bottom box in fig. 5). The advantage of using a uniform notation for constructions exhibiting different degrees of abstractness is that it allows for a unified bottom-up approach to research in CxG, thereby implementing a uniform notation for constructions with different levels of abstractness along the lexicon-syntax continuum.

6. Conclusions and outlook

In this paper I have discussed a number of problems with the interactions between lexical entries and grammatical constructions in Construction Grammar. Reviewing data on the ditransitive, causative-plus-*with*-adjunct, caused-motion, and resultative constructions as well as the [A *as* NP] pattern, I have argued (cf. Section 4.4) that postulating a strict division between lexical entries and grammatical

constructions is problematic because it cannot account for the full range of attested corpus data. More specifically, I have shown that efforts to specify abstract constraints limiting the interaction between lexical entries and grammatical constructions may ultimately lead to an increasing number of "theoretical" assumptions about the architecture of grammar. As the postulation of additional constraints may ultimately lead to theory-internally consistent but empirically unmotivated restrictions on the interactions between lexical entries and grammatical constructions (compare the discussion of the interaction between lexical entries and various modules and rule types in the Chomskyan tradition in Section 2), I have argued in this paper for a corpus-based bottom-up approach to CxG. The goal of this approach is to avoid problems arising from the interaction of lexical entries and abstract meaningful constructions by overcoming the artificial split between the lexicon and syntax that currently exists in Construction Grammar.

On this alternative view, Goldberg-style lexical entries are replaced with mini-constructions representing conventionalized form-meaning pairings of a verb's multiple senses. The inclusion of detailed information makes it possible to encode relevant semantic, pragmatic, and syntactic information whenever necessary, thereby achieving a greater level of precision. Using the same type of notation for more abstract and productive constructions, such as the *way*-construction or the subject-predicate construction, allows us to implement a unified representation of different entities along the lexicon-syntax continuum, thereby effectively eliminating the strict separation between the lexicon and syntax. The inclusion of frequency data in a construction's description makes it possible to investigate what types of correlations there are between a construction's frequency, its level of abstractness (semantic space), and its productivity (cf. Section 5.3).

Encoding more detailed information at different levels of the lexicon-syntax continuum means that independently existing meaningful constructions have a less powerful status than assumed by most CxG analyses. Their role is taken over primarily by information contained in conventionalized mini-constructions as well as analogical processes that produce non-conventionalized argument structures based on existing conventionalized form-meaning pairings in combination with contextual background information (see Boas, 2003ab). This step does not mean that independently existing meaningful constructions should be discarded. In Section 5.2.3 I have shown that the existence of Goldberg-type constructions is a natural by-product of the high type frequency of a given constructional pattern. That is, they represent a highly abstract schematization over a large number of conventionalized verb-senses instantiating a particular constructional pattern. While Goldberg-type constructions are important for the comprehension of novel utterances, they are not sufficient for predicting the full range of distribution of a verb's arguments in novel utterances based on non-conventionalized verb senses.

In other words, the production of novel utterances crucially relies on existing conventionalized mini-constructions, which encode detailed semantic, pragmatic, and syntactic information.

Clearly, future research in Construction Grammar must investigate the organization of the lexicon-syntax continuum in much greater detail. The Berkeley FrameNet project (Fillmore *et al.*, 2003, http://framenet.icsi.berkeley.edu) has already made some significant steps in this direction. It is currently expanding its frame-semantic analysis and description of the English lexicon to grammatical constructions. Among other things it seeks to compile a "constructicon" with corresponding annotated example sentences. Similar to lexical entries in FrameNet, the entries in the constructicon will (1) describe the constructions and their components, (2) set up construction elements (the syntactic elements that make up a construct), (3) explain the semantic contribution of the construction, (4) specify construction-to-construction relations, and (5) link construction descriptions with annotated sentences that exhibit their type (see Fillmore, 2008). Of particular importance will be the question of how to encode different types of semantic, pragmatic, and syntactic information in such a way that it is possible to account for a given utterance from a comprehension perspective, as well as a production perspective. The goals of the present paper have been more modest: to set out an approach for describing the lexicon-syntax continuum in a unified way, thereby overcoming problems currently surrounding the interaction of lexical entries and constructions in Construction Grammar.

Notes

* This paper is based on an invited plenary lecture delivered at the Second International Conference on Construction Grammar, organized by Jan-Ola Östman and Jaakko Leino at the University of Helsinki in October 2002. I would like to thank Jaakko Leino, Marc Pierce, and an anonymous reviewer for extensive comments on earlier versions of this paper.

1. The term "Chomskyan framework" refers to the different versions of Chomsky's syntactic theories such as *Aspects of the Theory of Syntax* (1965), *Lectures on Government and Binding* (1981), and *The Minimalist Program* (1995). All versions of the Chomskyan framework assume a strict separation between the lexicon and syntax.

2. See Webelhuth (1995, pp. 28–51) for an overview of different types of phrase structure rules.

3. In recent years, Jackendoff has been successful at overcoming the syntactico-centric methodology of most analyses couched in the Chomskyan framework. For example, Jackendoff (2002, pp. 148–149) proposes that syntax is by and large taken to be semantically motivated, while still allowing for some systematic residues of defective (i.e. autonomous) items and constructions. For details, see Goldberg & Jackendoff (2004) and Culicover & Jackendoff (2005).

4. There are a number of general constraints limiting the fusion of verbs with constructions, such as the Semantic Coherence Principle and the Correspondence Principle (Goldberg, 1995, p. 50). See Goldberg (1995, pp. 193–197) for a discussion of a number of constraints specific to the resultative construction. Note that for each of the constructions discussed by Goldberg (ditransitive, caused-motion, resultative, and *way*-constructions, among others), different sets of construction-specific constraints apply in order to restrict the application of these constructions. Imposing such restrictions is similar to the different types of constraints and filters used for various movement operations in the Chomskyan framework. That is, they serve to delimit the power of generative mechanisms in order to rule out unacceptable sentences.

5. Goldberg claims that in her approach the "mapping between semantics and syntax is done via constructions, not via lexical entries" (1995, p. 28). Certainly, this statement needs to be interpreted with caution because in order for a construction to be able to map from semantics to syntax, it first needs to access important information contained in a verb's lexical entry, such as information about semantic classes (cf. (14)–(16)) and collocational restrictions (cf. Tables 1 and 2). Given our discussion so far, it is clear that this information does indeed play a much more important role in the licensing of sentences than Goldberg currently suggests.

6. Goldberg's format for lexical entries is kept simple and only contains semantic information. For example, the lexical entry for the verb *wipe* specifies two participant roles, namely a 'wiper' role that is profiled and a 'wiped' role that is not profiled (Goldberg (1995, p. 189)). Although Goldberg admits that verbs are "associated with rich frame-semantic meanings" and that "richer aspects of verb meaning are required for aspects of linguistic theory" (1995, p. 29), she does not go into any detail as to how differences in verb meanings can be formalized more precisely. To this end, Nemoto (1998) and Boas (2003ab) point out that it is necessary to include much more detailed semantic information (as well as syntactic information when necessary) in a verb's lexical entry than what Goldberg actually encodes in her lexical entries.

7. Sufficient semantic and pragmatic overlap is present in cases in which a hearer can construe the novel meaning of the verb based on the existing conventionalized information associated with blow, sneeze, and contextual background information (see Boas, 2003a, pp. 264–278). This analogical process is constrained by Israel's (1996) Production Principle (Utterances should sound like things the speaker has heard before) and Comprehension Principle (Representations should capture similarities across experienced usages) as well as a number of other semantic and pragmatic restrictions. For more details, see Boas (2003a, pp. 260–277).

8. Note that there is a difference between production and comprehension of utterances (cf. Fillmore *et al.*'s (1988) difference between idioms of encoding and idioms of decoding). The production of sentences including novel verb senses (such as K*im sneezed the napkin off the table*) typically relies on a process of analogical association outlined above. In contrast, the comprehension of such novel verb senses may make use of more abstract constructions that are generalizations about the individual stored entities. Such a view is also compatible with Lehrer's (1990, p. 240) proposal regarding productivity. She points out that "because productivity is partial, the lexicon must contain the existing conventional senses as well as the rule."

9. Token frequency refers to the number of times a given instance (e.g., a particular word) is used in a particular construction. Type frequency refers to the number of distinct words that occur in a particular construction. (cf. Goldberg, 1995, p. 134)

10. Although such a model may be capable of predicting tendencies with respect to the language as a whole, it will still be difficult to predict specific choices made by a specific speech community or even an individual speaker in a particular situation. As such, the frequency information in Table 3 only represents the distribution found in the BNC and may differ between speakers and genres. This problem raises the question of how far the usage-based approach should be really applied to the description of language. Stefanowitsch & Gries (2003) argue for the inclusion of frequency data in the description of grammatical constructions. Their "collostructional approach" has the advantage that it increases "the adequacy of grammatical description by providing an objective way of identifying the meaning of a grammatical construction and determining the degree to which particular slots in it prefer or are restricted to a particular set of lexemes" (Stefanowitsch & Gries, 2003, p. 209).

11. Note that the two scales in Figure 4 do not always coincide. Thanks to Jaakko Leino for pointing this out.

12. In contrast to Goldberg (1995), who favors the inclusion of only semantic information in lexical entries (cf. her lexical entries (51a) and (51b) on p. 189), Boas (2003ab) proposes to include more detailed syntactic and pragmatic information. See Boas (2003a, pp. 159–213) for a formalization of mini-constructions (individual verb senses) including semantic, pragmatic, and syntactic information about their distributions.

References

Ackerman, F. & G. Webelhuth. (1998). *A Theory of Predicates*. Stanford: CSLI Publications.

Baker, C. (1999). *Seeing Clearly: Frame Semantic, Psycholinguistic, and Cross-Linguistic Approaches to the Semantics of the English verb 'See'*. Ph.D. Dissertation. University of California at Berkeley.

Baker, C., Fillmore, C.J. & Lowe, J.B. (1997). The Berkeley FrameNet Project. *COLING-ACL '98: Proceedings of the Conference*.

Barðdal, J. (to appear). *Productivity — Evidence from case and argument structure in Icelandic*. Amsterdam/Philadelphia: John Benjamins.

Bloomfield, L. (1933). *Language*. New York: Holt.

Boas, H.C. (2000). *Resultative Constructions in English and German.* Ph.D. Dissertation. University of North Carolina at Chapel Hill.

Boas, H.C. (2002a). On constructional polysemy and verbal polysemy in Construction Grammar. In V. Samiian (Ed.), *Proceedings of the 2000 Western Conference on Linguistics, 12* (pp. 126–139).

Boas, H.C. (2002b). Lexical or syntactic selection of non-subcategorized nominal arguments? The case of German *Sätzchen*. In I. Rauch & G. Carr (Eds.), *New Insights in Germanic Linguistics III* (pp. 9–26). New York: Peter Lang.

Boas, H.C. (2003a). *A Constructional Approach to Resultatives*. Stanford: CSLI Publications.

Boas, H.C. (2003b). Towards a lexical-constructional account of the Locative Alternation. In L., Carmichael, C.-H. Huang and V. Samiian (Eds.), *Proceedings of the 2001 Western Conference on Linguistics*, *13* (pp. 27–42).

Boas, H.C. (2005). Determining the productivity of resultative constructions: A reply to Goldberg & Jackendoff. *Language 81*(2), 448–464.

Boas, H.C. (2006). A frame-semantic approach to identifying syntactically relevant elements of meaning. In P. Steiner, H.C. Boas and S. Schierholz (Eds.), *Contrastive Studies and Valency. Studies in Honor of Hans Ulrich Boas* (pp. 119–149). Frankfurt/New York: Peter Lang.

Boas, H.C. (2007). Construction Grammar in the twenty-first century. *English Language and Linguistics 11*(3), 569–585.

Boas, H.C. (2008). Resolving form-meaning discrepancies in Construction Grammar. In J. Leino (Ed.), *Constructional Reorganization* (pp. 11–36). Amsterdam/Philadelphia: John Benjamins.

Boas, H.C. (to appear). Towards a frame-constructional approach of verb classification. *Revista Canaria de Estudios Ingleses, 57* (November 2008).

Bresnan, J. (2001). *Lexical-Functional Syntax*. Oxford: Blackwell.

Bybee, J. (1985). *Morphology. An Inquiry into the Relation between Meaning and Form*. Amsterdam/Philadelphia: John Benjamins.

Bybee, J. (2001). *Phonology and Language Use*. Cambridge: Cambridge University Press.

Chomsky, N. (1957). *Syntactic Structure*. The Hague: Mouton.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Croft, W. (2001). *Radical Construction Grammar*. Oxford: Oxford University Press.

Culicover, P. & R. Jackendoff. (2005). *Simpler Syntax*. Oxford: Oxford University Press.

DiSciullo, A.M. & E. Williams. (1987). *On the Definition of Word*. Cambridge, Mass.: MIT Press.

Croft, W. (2003). Lexical rules vs. constructions: a false dichotomy. In H. Cuyckens, T. Berg, R. Dirven, & K.-U. Panther (Eds.), *Motivation in Language: Studies in honour of Günter Radden* (pp. 49–68). Amsterdam/Philadelphia: John Benjamins.

Fillmore, C.J. (1977). Topics in Lexical Semantics. In P. Cole (Ed.), *Current Issues in Linguistic Theory* (pp. 76–136). Bloomington: Indiana University Press.

Fillmore, C.J. (1982). Frame Semantics. In Linguistic Society of Korea (Ed.), *Linguistics in the Morning Calm* (pp. 111–138). Seoul: Hanshin.

Fillmore, C.J. (1985). Frames and the semantics of understanding. *Quadernie di Semantica, 6*(2), 222–254.

Fillmore, C.J. (1988). The mechanisms of construction grammar. In S.A. Axmaker, A. Jaisser and H. Singmaster (Eds.), *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society* (pp. 35–55). University of California at Berkeley: Linguistics Department.

Fillmore, C.J. (2008). *Border Conflicts: FrameNet meets Construction Grammar*. Plenary Lecture at EURALEX XIII, Barcelona, Spain.

Fillmore, C.J., & B.T.S.Atkins. (2000). Describing polysemy: The case of *crawl*. In Y. Ravin & C. Laecock (Eds.), *Polysemy* (pp. 91–110). Oxford: Oxford University Press.

Fillmore, C.J., Johnson, C.R., & Petruck, M.R.L. (2003). Background to FrameNet. *International Journal of Lexicography, 16*(3), 235–250.

Fillmore, C.J., & P. Kay. (1993). *Construction Grammar*. Manuscript. University of California at Berkeley.

Fillmore, C.J., P. Kay and M.C. O'Connor. (1988). Regularity and idiomaticity in grammatical constructions: The case of 'let alone'. *Language, 64,* 501–538.

Fried, M. & J.-O. Östman. (2004). Construction Grammar. A thumbnail sketch. In M. Fried & J.-O. Östman (Eds.), *Construction Grammar in a Cross-Language Perspective* (pp. 11–86). Amsterdam/Philadelphia: Benjamins.

Goldberg, A. (1995). *Constructions: A Construction Grammar approach to argument structure*. Chicago: University of Chicago Press.
Goldberg, A. (1999). Construction Grammar. In K. Brown & J. Miller (Eds.), *Concise Encyclopedia of Grammatical Categories*. Amsterdam: Elsevier.
Goldberg, A. (2006). *Constructions at Work*. Oxford: Oxford University Press.
Goldberg, A. & R. Jackendoff. (2004). The English resultative as a family of constructions. *Language, 80*(3), 532–568.
Gries, S.T. (1999). Particle movement: A cognitive and functional approach. *Cognitive Linguistics, 10,*105–146.
Gries, S.T. & A. Stefanowitsch. (2003). Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics, 8*(2), 209–243,
Gross, M. (1979). On the failure of Generative Grammar. *Language,* 55, 859–858.
Gross, M. (1994). Constructing Lexicon Grammars. In B.T.S. Atkins & A. Zampolli (Eds.), *Computational Approaches to the Lexicon* (pp. 213–258). Oxford: Oxford University Press.
Harris, C. (1998). Psycholinguistic studies of entrenchment. In J.-P. Koenig (Ed.), *Conceptual Structure, Discourse and Language* (pp. 55–70). Stanford: CSLI Publications.
Hunston, S. & G. Francis. (2000). *Pattern Grammar. A Corpus-driven Approach to the Lexical Grammar of English*. Amsterdam/Philadelphia: Benjamins.
Israel, M. (1996). The way constructions grow. In A. Goldberg (Ed.), *Conceptual Structure, Discourse, and Language* (pp. 217–231). Stanford: CSLI Publications.
Iwata, S. (1998). *A Lexical Network Approach to Verbal Semantics*. Tokyo: Kaitakusha.
Iwata, S. (2002). Does MANNER count or not? Manner-of-motion verbs revisited. *Linguistics, 40*(1), 61–110.
Jackendoff, R. (1997a). *The Architecture of the Language Faculty*. Cambridge, MA: MIT Press.
Jackendoff, R. (1997b). Twistin' the night away. *Language, 73,* 534–559.
Jackendoff, R. (2002). *Foundations of Language: brain, meaning, grammar, evolution*. Oxford: Oxford University Press.
Kay, P. (1996). *Argument Structure: Causative ABC-Constructions*. Manuscript. University of California at Berkeley.
Kay, P. (2002a). An informal sketch of a formal architecture for Construction Grammar. *Grammars, 5,* 1–19.
Kay, P. (2002b). *Patterns of Coining*. Plenary Talk at the Second International Conference on Construction Grammar. University of Helsinki, October 2002.
Kay, P. & C.J. Fillmore. (1999). Grammatical constructions and linguistics generalizations: The *What's X doing Y?* construction. *Language, 75,* 1–33.
Kilgarriff, A. (1997). I don't believe in word senses. *Computers and the Humanities, 31*(2), 91–113.
Lakoff, G. (1987). *Women, Fire, and Dangerous Things*. Chicago: University of Chicago Press.
Langacker, R.W. (1987). *Foundations of Cognitive Grammar, Vol. I: Theoretical Prerequisites*. Stanford: Stanford University Press.
Langacker, R.W. (2000). A dynamic usage-based model. In M. Barlow & S. Kemmer (Eds.), *Usage Based Models of Language* (pp. 1–64). Stanford: CSLI Publications.
Lehrer, A. (1990). Polysemy, conventionality, and the structure of the lexicon. In *Cognitive Linguistics, 1*(2), 207–46.
Levin, B. (1993). *English Verb Classes and Alternations*. Chicago: University of Chicago Press.

Nemoto, N. (1998). On the polysemy of ditransitive *save*: the role of frame semantics in construction grammar. *English Linguistics, 15*, 219–242.
Petruck, M. (1996). Frame Semantics. In J. Verschueren, J.-O. Östman, J. Blommaert, & C. Bulcaen (Eds.), *Handbook of Pragmatics* (pp. 1–13). Amsterdam and Philadelphia: Benjamins.
Pinker, S. (1989). *Learnability and Cognition*. Cambridge, MA: MIT Press.
Pollard, C. & I. Sag. (1994). *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.
Rappaport Hovav, M. & B. Levin. (1998). Building verb meaning. In M. Butt & W. Geuder (Eds.), *The Projection of Arguments* (pp. 97–134). Stanford: CSLI Publications.
Rappaport Hovav, M. & B. Levin. (2001). An event structure account of English resultatives. *Language, 77*, 766–797.
Ruiz de Mendoza Ibáñez, F.J. & R. Mairal. (2007). High-level metaphor and metonymy in meaning constructions. In G. Radden, K.-M. Köpcke, T. Berg, and P. Siemund (Eds), *Aspects of Meaning Construction in Lexicon and Grammar* (pp. 44–49). Amsterdam/Philadelphia: John Benjamins.
Ruiz de Mendoza Ibáñez, F.J. & R. Mairal. (2008). Levels of description and constraining factors in meaning construction: An introduction to the Lexical Constructional Model. *Folia Linguistica, 42*(3/4), forthcoming.
Sampson, G. (2001). *Empirical Linguistics*. London/New York: Continuum International.
Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
Stefanowitsch, A. & S.T. Gries. (2003). Collostructions: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics, 8*(2), 209–243.
Van Riemsdijk, H. & E. Williams. (1986). *Introduction to the Theory of Grammar*. Cambridge, Mass.: MIT Press.
Verspoor, C.M. (1997). *Contextually-Dependent Lexical Semantics*. Ph.D. Dissertation, University of Edinburgh.
Webelhuth, G. (1995). X-bar Theory and Case Theory. In G. Webelhuth (Ed.), *Government and Binding Theory and the Minimalist Program* (pp. 15–86). Oxford: Blackwell.

Author's address

Hans C. Boas
Associate Professor for Germanic Linguistics
Department of Germanic Studies
1 University Station C 3300, EPS 3.102
The University of Texas at Austin
Austin, TX 78712–0304, U.S.A.
Phone: (512) 471–4123
Fax: (512) 471–4025

hcb@mail.utexas.edu
http://hcb.gmc.utexas.edu

About the author

Hans C. Boas is an Associate Professor of Germanic Linguistics at The University of Texas at Austin. Before moving to Texas he received his undergraduate education in linguistics and law

at the University of Göttingen (Germany), and his M.A. (1995) and Ph.D. (2000) in linguistics at the University of North Carolina at Chapel Hill. Funded by the German Academic Exchange Service, he was a postdoctoral researcher with the FrameNet project at the International Computer Science Institute in Berkeley (2000–2001). His primary research interests are Construction Grammar, Frame Semantics, Computational Lexicography, Documentary Linguistics, Language Variation, and Language Contact and Language Death. Since 2001, he has directed the Texas German Dialect Project (http://www.tgdp.org), which documents the remnants of Texas German, an immigrant dialect spoken for more than 150 years and now critically endangered. For his work on Texas German he received a fellowship from the National Endowment for the Humanities (2005) as well as the Hugo-Moser Prize for Germanic Linguistics from the Institute for the German language in Mannheim (Germany) (2007).